On the Timing and Pricing of Dividends

Jules H. van Binsbergen†
Stanford GSB and NBER

Michael W. Brandt‡
Fuqua School of Business and NBER

Ralph S.J. Koijen§
Chicago Booth and NBER

This version: August 2010

Abstract

We use data from options and futures markets to recover the prices of dividend strips on the aggregate stock market. The price of a \( k \)-year dividend strip is the present value of the dividend paid in \( k \) years. The value of the stock market is the sum of all dividend strip prices across maturities. We study the asset pricing properties of strips and find that expected returns, Sharpe ratios, and volatilities on short-term dividend strips are higher than on the aggregate stock market, while their CAPM betas are well below one. Prices of short-term dividend strips are more volatile than their realizations, leading to excess volatility and return predictability. The properties of dividend strip prices implied by leading asset pricing models differ from those observed in the data.

First draft: June 2009. We are grateful to Ravi Bansal, Robert Battalio, Alessandro Beber, John Campbell, John Cochrane, George Constantinides, Zhi Da, Peter DeMarzo, Joost Driessen, Darrell Duffie, Pengjie Gao, John Heaton, Xavier Gabaix, Nicolae Garleanu, Lars Hansen, Otto van Hemert, Mark Hendricks, Dirk Jenter, Martin Lettau, Lars Lochstoer, Hanno Lustig, Pascal Maenhout, Toby Moskowitz, Christian Mueller-Glissmann, Stefan Nagel, Stavros Panageas, Lubos Pastor, Lasse Pedersen, Anamaria Pieschacon, Rob Stambaugh, Sheridan Titman, Stijn Van Nieuwerburgh, Adrien Verdelhan, Pietro Veronesi, Rob Vishny, Evert Vrugt, Jessica Wachter, Jeff Wurgler, Amir Yaron, Motohiro Yogo and seminar participants APG Investments, AQR Capital, Berkeley Haas, Boston College, Brigham Young University, CEPR Gerzensee meetings, Chicago Booth, NBER Asset Pricing Meetings, Notre Dame, Stanford GSB, Tilburg University, UT Austin, and the Empirical Asset Pricing Retreat at the University of Amsterdam for useful comments and discussions. We thank CRSP for generous financial support.

†jvb2@gsb.stanford.edu, (650)-721-1353
‡mbrandt@duke.edu, (919)-660-1948
§ralph.koijen@chicagobooth.edu, (773)-834-4199.
A central question in asset pricing is how to discount future cash flows to obtain today’s value of an asset. For instance, total wealth is the price of a claim to all future consumption (Lucas (1978)). Similarly, the value of the aggregate stock market equals the sum of discounted future dividend payments (Gordon (1962)). The majority of the equity market literature has focused on the dynamics of the value of the aggregate stock market. However, in addition to studying the value of the sum of discounted dividends, exploring the properties of the individual terms in the sum, also called dividend strips, provides us with a lot of information about the way stock prices are formed. Analogously to zero-coupon bonds, which contain information about discount rates at different horizons for fixed income securities, having information on dividend strips informs us about discount rates of risky cash flows. Studying dividend strips can therefore improve our understanding of investors’ risk preferences and the endowment or technology process in macro-economic models. This paper is the first to empirically measure the prices of dividend strips. Our approach only requires no-arbitrage relations and does not rely on a specific model.

In this paper, we decompose the S&P 500 index, which is a broad US equity index, into a portfolio of short-term dividend strips, which we call the short-term asset, and a portfolio of long-term dividend strips, which we call the long-term asset. The short-term assets entitles the holder to the realized dividends of the index for a period of up to three years. Our main focus is to compare the asset pricing properties of the short-term asset to those of the index, both empirically and theoretically.

More formally, the value of equity index $S_t$ is given by the discounted value of its dividends $(D_{t+i})_{i=1}^{\infty}$:

$$S_t = \sum_{i=1}^{\infty} E_t (M_{t:t+i} D_{t+i}),$$

with $M_{t:t+i} = \prod_{j=1}^{i} M_{t+j}$, the product of stochastic discount factors. We can decompose the stock index as:

$$S_t = \underbrace{\sum_{i=1}^{T} E_t (M_{t:t+i} D_{t+i})}_{\text{price of the short-term asset}} + \underbrace{\sum_{i=T+1}^{\infty} E_t (M_{t:t+i} D_{t+i})}_{\text{price of the long-term asset}},$$

where the short-term asset is the price of all dividends up until time $T$, and the long-term asset is the price of the remaining dividends. To compute the price of the short-term asset, we use a newly-constructed data set on options and futures on the S&P500 index.
Using this method, we document five properties of the short-term asset in comparison to the aggregate stock market. First, expected returns, volatilities, and Sharpe ratios on the short-term asset are on average higher. Second, the CAPM beta of the short-term asset returns with respect to the aggregate stock market is 0.5. Third, the CAPM alpha of short-term asset returns is 10% per year. Fourth, the prices of the short-term asset are more volatile than their realizations, pointing to excess volatility. Fifth, the returns on the short-term asset are strongly predictable.

To provide a theoretical benchmark for our results, we compute dividend strips in several leading asset pricing models. Recent consumption-based asset pricing models have made substantial progress in explaining many asset pricing puzzles across various markets. Even though such models are not often used to study the pricing of dividend strips, they do have theoretical predictions about their values, which we explore in this paper.\footnote{Notable exceptions are Lettau and Wachter (2007) and Lettau (2009).} We focus on the external habit formation model of Campbell and Cochrane (1999), the long-run risks model of Bansal and Yaron (2004), and the variable rare disasters model of Gabaix (2009). We find that both the long-run risks model and the external habit formation model predict that expected returns, volatilities, and Sharpe ratios of short-term dividend strips are near zero and lower than those of the aggregate market. In the rare disasters model, the volatilities and Sharpe ratios of short-term dividend strips are lower than the aggregate market. Expected returns on the other hand are equal across all maturities of dividend strips, and therefore also equal to those on the aggregate market. Our results point to the importance of short-term risk compensation equity markets, and this compensation is higher than predicted by leading asset pricing models.

Our results have several important additional implications for empirical and theoretical asset pricing. First, since Shiller (1981) pointed out that stock prices are more volatile than subsequent dividend realizations, the interpretation has been that discount rates fluctuate over time and are persistent. The long duration of equity makes prices very sensitive to small persistent movements in discount rates, thereby giving rise to excess volatility. We show, however, that the same phenomenon arises for the short-term asset. This suggests that a complete explanation of excess volatility is able to generate excess volatility both for the aggregate stock market and for short-term dividends. The excess variation in prices also suggests that discount rates fluctuate, and we should therefore find that prices, normalized by some measure of dividends, forecast returns on the short-term asset. We show that this is indeed the case, leading to the sixth property. Second, the first four properties we document, combined with the fact that the CAPM alphas are...
virtually unaffected if we include additional well-known asset pricing factors such as size, value or momentum, suggest that short-term assets are potentially important new test assets that might be useful in cross-sectional asset pricing tests.  

Our paper relates to Lettau and Wachter (2007) and Croce, Lettau, and Ludvigson (2009). Lettau and Wachter (2007) argue that habit formation models as in Campbell and Cochrane (1999), generate higher expected returns for long-term dividend strips as shocks to the discount factor are priced. Firms with long-duration cash flows have a high exposure to such shocks, and should therefore have a higher risk premium than firms with short-duration cash flows. If one adheres to the view that value firms have short-duration cash flows and growth firms have long-duration cash flows, this would imply that there is a growth premium, not a value premium. Lettau and Wachter (2007) propose a reduced-form model that generates higher expected returns for short-term dividend strips. They illustrate the correlation structure between (un)expected cash flow shocks and shocks to the price of risk and stochastic discount factor that is sufficient to generate a value premium in their model. Croce, Lettau, and Ludvigson (2009) argue that the long run risk model as proposed by Bansal and Yaron (2004) also generates higher risk premia for long-term dividend strips. However, Croce, Lettau, and Ludvigson (2009) also show that if the agent in the model cannot distinguish between short-term and long-term shocks, risk premia on short-term dividend strips can be higher.

Studying the properties of short-term assets is not only interesting from an academic perspective. Recently, dividend strips, futures, and swaps have received a lot of attention in the practitioners’ literature. Several banks are offering dividend swaps on a range of stock indices. With such a contract, the dividend purchaser pays the market-implied level that is derived from an equity index derivative multiplied by the overall exposure per index point (the fixed leg). The counterparty, with a long position in the equity index, pays the realized dividend level multiplied by the exposure per index point. For the S&P500, Standard and Poor’s has introduced the S&P500 Dividend Index, which is a running total of dividend points. The index is reset to zero after the close on the third Friday of the last month of every calendar quarter, to coincide with futures and options expirations. It measures the total dividend points of the S&P500 since the previous reset date and is used by derivative traders to hedge their dividend positions. Further, from 1982 to 1992, investors could invest in derivatives at the American Stock Exchange (AMEX) that split the total return on individual stocks into a price appreciation part and a dividend

2Lewellen, Nagel, and Shanken (2009) argue that the standard set of test assets has a strong factor structure, and that it would be valuable to have a new set of test assets.

yield part. Also in the UK, split-capital funds offered financial instruments that separate investment in a fund’s price appreciation and its dividend stream in the late 90s. Wilkens and Wimschulte (2009) discuss the European market of dividend futures that started mid-2008.

We proceed as follows. In Section 1 we discuss various ways to trade the short-term asset, either by creating it synthetically using index derivatives or by trading dividend derivatives directly. Section 3 discusses our dataset. In Section 3 we discuss the empirical results. In Section 4, we compare our findings with several leading asset pricing models. We present several robustness checks in Section 5. In Section 6 we discuss several possible extensions. Section 7 concludes.

1 The market for dividends

There are two ways to trade dividends in financial markets. First, dividend strips can be replicated using options and futures data, which is the approach we follow in this paper. In 1990, the Chicago Board Options Exchange (CBOE) introduced Long-Term Equity Anticipation Securities (LEAPS) options. The owner of a call (put) option has the right to purchase (sell) the stock index at maturity at a predetermined price $X$. LEAPS have maturities up to three years. The maximum maturity of LEAPS for the sample period in our dataset is displayed in Figure 1. As a result of the issuing cycle of LEAPS, the maximum available maturity changes over time. The set of maturities of these claims is not constant and varies depending on the issuing cycle. On average, there are around six maturities greater than three months available at any particular time.

To compute dividend strip prices from options data, we only require the absence arbitrage opportunities. Under this condition, the put-call parity for European options holds (Stoll (1969)):

$$c_{t,T} + X e^{-r_{t,T}(T-t)} = p_{t,T} + S_t - P_{t,T}, \quad (1)$$

where $p_{t,T}$ and $c_{t,T}$ are the prices of a European put and call option at time $t$, with maturity $T$, and strike price $X$. $r_{t,T}$ is the interest rate between time $t$ and $T$. We use the symbol $P_{t,T}$ to denote the value of the short-term asset, which we defined in the introduction as:

$$P_{t,T} \equiv \sum_{i=1}^{T} E_t (M_{t+i}D_{t+i}). \quad (2)$$
We can rewrite (1) to obtain the price of the dividend strip:

$$P_{t,T} = p_{t,T} - c_{t,T} + S_t - X e^{-r_{t,T}(T-t)}. \quad (3)$$

This parity relation shows that purchasing a dividend strip is equivalent to buying a put option, writing a call option, buying the stocks in the index, and borrowing cash.

A second way to synthetically create the short-term asset is by using futures contracts. The owner of the futures contract agrees to purchase the stock index for a predetermined price, $F_{t,T}$. Absence of arbitrage opportunities implies the cost-of-carry formula for equity futures:

$$P_{t,T} = S_t - e^{-r_{t,T}(T-t)} F_{t,T}. \quad (4)$$

Hence, buying a dividend strip is the same as buying the stock index and selling a position in a futures contract. In both cases, we exploit that payoffs of derivatives contracts are based on the ex-dividend price, which allows us to recover the price of the short-term asset.

In addition to computing the prices of short-term asset using equity derivatives, it is also possible to trade dividends directly via dividend derivatives such as dividend swaps, dividend futures, and options on dividends. Most transactions take place in over-the-counter (OTC) markets, but several exchange-traded products have been introduced recently. For instance, the CBOE introduced options on S&P500 index dividends in May 2010. This development follows the introduction of an array of dividend derivatives at the Eurex. The Eurex introduced on June 30, 2008 dividend futures on the Dow Jones EURO STOXX 50 Index and in February 2010, futures are available on five different indices. In addition, the Eurex now introduced dividend futures on the constituents of the Dow Jones EURO STOXX 50 in January 2010. As measured by open interest, the size of the market for index dividend futures is already 20% of the size of the market for index futures at March 30 2010, illustrating the rapid developments of dividend trading. The advantage of dividend derivatives is that the maturities available are longer and up to 15 years.

---


6The dollar volume averages to $33.4 Billion in 2009, see: http://www.reuters.com/article/idUSLDE60A1DO20100111.
2 Dividend prices in a Lucas economy

To illustrate how dividend prices can be understood in an equilibrium model, we compute the dividend strip prices in the consumption CAPM of Lucas (1978). In Section 3, we extend these results to more recent consumption-based asset pricing models. Consumption growth, \( \Delta c_t = \log C_t - \log C_{t-1} \), is assumed to be i.i.d.:

\[
\Delta c_{t+1} = g + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N(0, \sigma^2),
\]

where \( g \) is the constant average growth rate and \( \sigma \) the growth rate volatility. The representative agent has time-separable CRRA utility:

\[
E_t \left( \sum_{s=0}^{\infty} \beta^s \frac{C_{t+s}^{1-\gamma}}{1-\gamma} \right).
\]

The one-period stochastic discount factor is in this case given by:

\[
M_{t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma},
\]

and the \( k \)-period stochastic discount factor equals:

\[
M_{t:t+k} = \beta^k \left( \frac{C_{t+k}}{C_t} \right)^{-\gamma}.
\]

Assuming that the aggregate dividend equals aggregate consumption, \( C_t = D_t \), the price of the \( k \)-period short-term asset is given by:

\[
P_{t,t+k} = \sum_{s=1}^{k} \exp \left( s \log \beta + s(1-\gamma)g + \frac{s}{2}(1-\gamma)^2 \sigma^2 \right).
\]

3 Data and dividend strategies

3.1 Data sources

We measure dividend prices using put-call parity, which is a no-arbitrage relationship, in equation (1). To compute dividend prices as accurately as possible, we record each of the components in equation (1) within the same minute of the last trading day of each month. To this end, we use data from four different sources. First, we use a new data set provided by the CBOE containing intra-day trades and quotes on S&P500 options.
between January 1996 and May 2009. The data contains information about all option contracts for which the S&P500 index is the underlying asset. Second, we obtain minute-level data between January 1996 and May 2009 of the index values and futures prices of the S&P500 index from Tick Data Inc. Third, the interest rate is calculated from a collection of continuously-compounded zero-coupon interest rates at various maturities and provided by IvyDB (OptionMetrics). This zero curve is derived from LIBOR rates from the British Bankers’ Association (BBA) and settlement prices of Chicago Mercantile Exchange (CME) Eurodollar futures. For a given option, the appropriate interest rate input corresponds to the zero-coupon rate that has a maturity equal to the option’s expiration date, and is obtained by linearly interpolating between the two closest zero-coupon rates on the zero curve. Fourth, to compute daily dividends, we obtain daily return data with and without distributions from S&P index services. Cash dividends are then computed as the difference between these two returns, multiplied by the lagged value of the index.

### 3.2 Data selection and matching

As mentioned before, our data allows us to match call and put option prices and index values within a minute interval. We therefore select option quotes for puts and calls between 10am and 2pm that trade within the same minute, and match these quotes with the tick-level index data, again within the minute. Changing the time interval to either 10am to 11am or 1pm to 2pm has no effect on our results, as we demonstrate in Section 6.

We compute dividend prices at the last trading day of the month. For a given strike price and maturity, we collect all quotes on call option contracts and find a quote on a put option contract, with the same strike price and maturity, that is quoted closest in calendar time. Of the resulting matches, we keep the matches for each strike and maturity that trade closest to each other in time. This typically results in a large set of matches for which the quotes are recorded within the same second of the day, making the matching procedure as precise as possible. For each of these matches, we use the put-call parity

---

7 We use data from Bloomberg to replicate the OptionMetrics yield curves and obtain very similar results.

8 Alternative interpolation schemes give the same results at the reported precision.

9 Using closing prices from OptionMetrics for all quantities does not guarantee that the index value and option prices are recorded at the same time and induces substantial noise in our computations, see also Constantinides, Jackwerth, and Perrakis (2009)). For instance, the options exchange closes 15 minutes later than the equity exchange, which leads to much wider bid-ask spreads in options markets during this period. OptionMetrics reports the last quote of the trading day, which is likely to fall in this 15-minute interval. We reproduced our results using OptionMetrics data, and find similar results for average returns, but the volatility of prices and returns is substantially higher.
relation to calculate the price of the dividend strip. We then take the median across all prices for a given maturity, resulting in the final price we use in our analysis. By taking the median across a large set of dividend prices, we mitigate potential issues related to measurement error or market microstructure noise.

To illustrate the number of matches we find for quotes within the same second, Figure 2 reports the average number of quotes per maturity during the last trading day of the month in a particular year. We focus on option contracts with a maturity between 1 and 2 years. The number of quotes increases substantially over time, presumably as a result of the introduction of electronic trading. However, even in the first year of our sample, we have on average nearly a thousand matches per maturity on a given trading day for options with maturities between 1 and 2 years.

### 3.3 Dividend strategies

Holding a long position in dividends has the potential disadvantage that a long position in the index is required (see equations (11) and (14)). As index replication is not costless, we also consider investing in a so-called dividend steepener. This asset entitles the holder to the dividends paid out between period $T_1$ and $T_2$, $T_1 < T_2$. The price of the dividend steepener is given by:

$$ P_{t,T_1,T_2} = P_{t,T_2} - P_{t,T_1} = p_{t,T_2} - p_{t,T_1} - c_{t,T_1} - X \left( e^{-r_{t,T_2}(T_2-t)} - e^{-r_{t,T_1}(T_1-t)} \right). \quad (10) $$

This strategy can be interpreted as buying $T_2$ periods of dividends and selling the first $T_1$ periods of dividends. This strategy does not involve any dividend payments until time $T_1$. Replicating this asset does not require a long position in the index and simply involves buying and writing two calls and two puts, in addition to a cash position. The dividend steepener is also interesting to study as a macro-economic trading strategy, as it can be used to bet on the timing of a recovery of the economy following a recession. During severe recessions, firms slash dividends and increase them when the economy rebounds. By choosing $T_1$ further into the future, investors bet on a later recovery.\footnote{See also “Dividend Swaps Offer Way to Pounce on a Rebound,” Wall Street Journal, April 2009.}

By applying the cost-of-carry formula for equity index futures to two different maturities, the price of the dividend steepener for period $T_1$ to $T_2$, $T_1 < T_2$, can also
be computed using futures contracts as:

\[ \mathcal{P}_{t,T_1,T_2} = e^{r_t(T_1-t)}F_{t,T_1} - e^{r_t(T_2-t)}F_{t,T_2}. \]

In this case, the steepener only involves two futures contracts and does not require any trading of the constituents of the index. By no-arbitrage, the prices implied by equity options and futures need to coincide. Since LEAPS have longer maturities than index futures, we rely on options for most of our analysis. For the maturities for which both futures and options data is available, we show below that the prices obtained from either market are close, and our main findings are unaffected by using either options or futures.

Apart from reporting dividend prices, we also implement two simple trading strategies. The first trading strategy goes long in the short-term asset. The monthly return series on this strategy is given by:

\[ R_{1,t+1} = \frac{\mathcal{P}_{t+1,T_1-1} + D_{t+1}}{\mathcal{P}_{t,T}}. \]

The monthly returns series on the dividend steepener is given by:

\[ R_{2,t+1} = \frac{\mathcal{P}_{t+1,T_1-1,T_2-1}}{\mathcal{P}_{t,T_1,T_2}}, \]

which illustrates that this return strategy does not return any dividend payments until time \( T_1 \). Further details on the implementation of these strategies can be found in Appendix A.

4 Main empirical results

In this section, we document the properties of the prices and returns on the short-term asset. First, we study dividend prices in Section 4.1. In Section 4.2 we study the properties of dividend returns. In the remaining subsections, we study excess volatility of dividend strip prices, and the predictability of the return series that we compute.

4.1 Properties of dividend prices

Figure 3 displays the prices of the first 6, 12, 18, and 24 months of dividends during our sample period. To obtain dividend prices at constant maturities, we interpolate over the available maturities. For instance, in January 1996, the price of the dividends paid out between that date and June 1997 is $20. As expected, the prices monotonically
increase with maturity. Violations of this condition would imply the existence of arbitrage opportunities. Further, the dividend prices for all maturities, drop during the two NBER recessions in our sample period, which occur between March and November 2001 and the recent recession. This is to be expected, as during recessions expected growth of dividends drops and discount rates are likely to increase. This effect is more pronounced for the 24-month price. The 6-month price is less volatile.

As dividend prices are non-stationary over time, it is perhaps more insightful to scale dividend prices by the value of the S&P500 index. In Figure 4 we plot the prices of the first 6, 12, 18, and 24-month dividend prices as a fraction of the index value. The ratios are highly correlated. They drop between 1997 and 2001, and slowly increase afterwards. When comparing Figure 4 with Figure 3 one interesting observation is that during the recession of 2001, both the ratio and the level of dividend prices drop, whereas for the recent recession the level of dividend prices drops, but not by as much as the index level. This leads to an increase in the ratio. One interpretation of this finding is that the most recent recession has a longer-lasting impact than the recession in 2001. The index level is more sensitive to revisions in long-term cash flow expectations and discount rates than the short-term asset, see Shiller (1981) and Lettau and Wachter (2007). A more severe recession can therefore lead to a larger decline in the index value than in the price of the short-term asset.

4.2 Properties of dividend returns

We now report the return characteristics of the two investment strategies. Figure 5 and Figure 6 plot the time series of monthly returns on the two trading strategies. Figure 7 and Figure 8 display the histogram of returns. The two trading strategies are highly positively correlated, with a correlation coefficient of 92.2%. Panel A of Table 1 lists the summary statistics alongside the same statistics for the S&P500 and the market for the full sample period. For the market, we use the CRSP value-weighted return of all stocks traded on the AMEX, NYSE, and Nasdaq. Both dividend strategies have a high monthly average return equal to 1.20% (annualized 14.4%) for trading strategy 1 and 1.15% (annualized 13.8%) for trading strategy 2. Over the same period, the average return on the market portfolio was 0.54% (annualized 6.5%) and the return on the S&P500 index was only 0.49% (annualized 5.9%). The higher average returns also come with a higher level of volatility than both the aggregate stock market and the S&P500 index, with monthly return volatilities of 7.9% for strategy 1 and 9.8% for strategy 2. Over the same period the monthly market volatility is 4.9% and the volatility of the return on the S&P500
index equals 4.7%. The summary statistics also indicate that dividend returns tend to have fatter tails than both equity indices. Despite the higher volatility, the dividend strategies result in substantially higher Sharpe ratios. The Sharpe ratios of the dividend strategies are about twice as high as the Sharpe ratios of both the aggregate stock market and the S&P500 index. Duffee (2010) shows that Sharpe ratios are lower for Treasury bonds with longer maturities. We document a similar property in equity markets; Sharpe ratios are lower for dividend claims with longer maturities.

We find that the volatility of dividend returns is lower in the second part of our sample. To further analyze the volatility of dividend returns, we estimate a GARCH(1,1) model for each return series and for the returns on the aggregate stock market. In Figure 10, we show that the volatility of dividend returns and the aggregate stock market broadly follow the same pattern. The correlation between the volatility of the dividend returns of strategy 1 and the S&P500 is 0.52. Table 4 reports the estimates of the GARCH(1,1)-specification, illustrating that the parameters of the volatility equations are very similar as well.

The volatility of the two return strategies was substantially higher before 2003 than after. Table 1 therefore also presents summary statistics for the period before January 2003 (Panel B) and for the period afterwards (Panel C). We are mostly interested in the average return and volatility of the dividend strategies relative to the same statistics of the equity indices. We find consistently across both sample periods that the average return on the dividend strategies is about twice as high as the average return on the market or the S&P500 index. The volatility of the dividend strategies is high in both sub-periods, even though the volatilities in the more recent sample are much closer to the levels of volatility that we record for the equity indices. The Sharpe ratios of the dividend strategies are comparable across subperiods, and always substantially higher than the ones of either the aggregate stock market or the S&P500 index. Overall, the conclusions we draw from the full sample are consistent with our findings in both sub-samples.

Table 2 presents OLS regressions of the two trading strategies on the market portfolio in excess of the one-month short rate. We find that both dividend strategies have a CAPM beta of 0.49. Secondly, $R^2$ values of the regression are low and close to 10%. The intercept of the regression equals 0.79% for strategy 1 and 0.73% for strategy 2, which in annualized terms equals 9.48% and 8.76%. Despite these economically significant intercepts, the results are not statistically significant at conventional levels due to the

\footnote{Using S&P500 index returns instead of the market portfolio leads to almost identical regression results.}
substantial volatility of these two return strategies and the rather short time series that is available for dividend returns. Generally the p-values are just above 10%.

Table 3 presents regression results for the Fama-French three factor model, in which we also include the AR(1)-term in the second and fourth column. The market beta is unaffected by the additional factors and is estimated between 0.5 and 0.6, depending on the strategy and specification. We find positive loadings on the book-to-market factor, which seems consistent with duration-based explanations of the value premium. The loading is statistically significant in case of the dividend steepener. The coefficient on the size portfolio switches sign depending on the specification, and has very low significance. Perhaps most interestingly, the intercepts are hardly affected by including additional factors; monthly alphas are estimated between 0.56% and 0.70%. These results suggest that the short-term asset has rather high expected returns that cannot be explained easily by standard asset pricing models.

The high monthly alphas compensate investors for the risk in the dividend strategy that cannot be explained by the other priced factors. Our results becomes even more striking, however, if we account for the fact that dividend growth rates are, to some extent, predictable, see for instance Lettau and Ludvigson (2005), Ang and Bekaert (2006), Chen, Da, and Priestley (2009), and Binsbergen and Koijen (2010). To illustrate the degree of dividend growth predictability in the S&P500 during various sample periods, we follow the approach developed in Binsbergen and Koijen (2010) to obtain an estimate of expected dividend growth rates. They combine standard filtering techniques with a present-value model as in Campbell and Shiller (1988) to forecast future returns and dividend growth rates. The approach is summarized in Appendix B.

The estimation results are summarized in Table 5. We provide parameter estimates for three data periods, the post-war period, starting in 1946, the period for which monthly data on the index is available, starting in 1970, and the period for which daily data is available, starting in 1989. Consistent with Binsbergen and Koijen (2010), we find that both expected returns and expected dividend growth rates are predictable. Further, both expected returns and expected dividend growth rates have a persistent component, but expected returns are more persistent than expected dividend growth rates. Interestingly, as the starting date of our sample period increases, both the $R^2$ value of returns as well as the $R^2$ value of dividend growth rates strongly increases. Over the data period starting in 1989, we find an $R^2$ value for dividend growth rates of 56%.

This high level of dividend growth predictability combined with the high volatility of the returns on the short-term dividend claim seems rather puzzling. The volatility
of annual dividend growth rates is only 7%, but a substantial part of the variance can be explained by simply predictor variables. As such, it would seem that to correctly price claims on the S&P500 index, we need a model that generates a downward sloping term structure of expected returns and volatilities, and which generates, or allows for, a non-trivial degree of dividend growth predictability. The unpredictable part of dividend growth, which is rather small, then needs to be highly priced.

4.3 Excess volatility of short-term dividend claims

Shiller (1981) points out that prices are more volatile than subsequent dividends, which is commonly known as “excess volatility.” One explanation has been that discount rates fluctuate over time and are persistent. The long duration of equity makes prices very sensitive to small movements in discount rates, thereby giving rise to excess volatility.

Since we study short-term claims, we can directly compare prices to subsequent realizations. Figure 9 plots the price of the next year of dividends and the realized dividends during the next year. We shift the latter time series such that the price and subsequent realization are plotted at the same date to simplify the comparison. This illustrates that the high volatility of dividend returns is mostly coming from variation in dividend prices as opposed to their realizations. This points to “excess volatility” at the short end of the equity curve. An explanation of the excess volatility puzzle therefore ideally accounts for both the excess volatility of the equity index as well as that of the short-term assets.

4.4 Predictability of dividend returns

The previous section shows that prices are more variable than subsequent realizations. This suggests that discount rates fluctuate over time, which in turn implies that we need to be able to uncover a predictable component in the returns on dividend strategies. Some of this evidence is present already in Table 2, which shows that dividend returns are to a certain extent mean-reverting. We extend this evidence by regressing monthly dividend returns on the lagged value of $P_{t,T}/D_t$. This is the equivalent of the price-dividend ratio for the short-term asset.\footnote{See, among others, Fama and French (1988), Campbell and Shiller (1988), Cochrane (1991), Cochrane (2006), Lettau and Van Nieuwerburgh (2006), Wachter and Warusawitharana (2009), and Binsbergen and Koijen (2010) for the predictability of returns by the dividend yield.} The results are presented in Table 6. We find that $P_{t,T}/D_t$ forecasts dividend returns with a negative sign, and is highly significant. We use OLS standard errors to determine the statistical significance of the predictive
coefficient. To mitigate concerns regarding measurement error in the predictor variable, we smooth $P_{t,T}/D_t$ over three months and use this predictor variable instead. The results are reported in the second column and are very comparable to the first column.

5 Comparison with asset pricing models

To provide a theoretical benchmark for our results, we compute dividend strips in several leading asset pricing models in this section. Recent consumption-based asset pricing models have made substantial progress in explaining many asset pricing puzzles across various markets. Even though such models are not often used to study the pricing of dividend strips, they do have theoretical predictions about their values. We consider the Campbell and Cochrane (1999) external habit formation model, the Bansal and Yaron (2004) long-run risk model, the Barro-Rietz rare disasters framework (Barro (2006)) as explored by Gabaix (2009) and Wachter (2010). We focus on the calibration of Gabaix (2009) in this case.

The habit model and the long-run risk model imply that the risk premium and volatility on long-term dividend claims are higher. The risk premium on the short-term asset is virtually zero and lower than on the aggregate stock market, which is contrary to what we measure in the data. In the rare disasters model, expected returns are constant across maturities, but the volatilities are higher for long-term dividend claims than for short-term claims. To generate these results, we use the original calibrations that are successful in matching facts about the aggregate stock market. It is important to keep in mind though that such models have a relatively simple shock structure and have not been calibrated to match prices of dividend strips. It may be possible to consider alternative calibrations or model extensions that do match the features of dividend strip prices we report.

We also consider the model of Lettau and Wachter (2007) who exogenously specify the joint dynamics of cash flows and the stochastic discount factor to match the value premium. In their model, expected returns and volatilities of the short-term asset are higher than on the aggregate stock market, and the CAPM beta of the short-term asset is well below one, resulting in a substantial CAPM alpha. These features of their model are in line with our empirical findings.

\footnote{See Cochrane and Piazzesi (2005) for a similar treatment of measurement error in the forecasting variable of, in their case, bond returns.}
5.1 External habit formation

We first summarize the key equations of the Campbell and Cochrane (1999) habit formation model. The stochastic discount factor is given by:

\[ M_{t+1} = \delta G^{e^{-\gamma(s_{t+1}-s_t+v_{t+1})}}, \]  

(13)

where \( G \) represents consumption growth, \( \gamma \) is the curvature parameter, \( v_{t+1} \) is unexpected consumption growth, and \( s_t \) is the log consumption surplus ratio whose dynamics are given by:

\[ s_{t+1} = (1 - \phi)\bar{s} + \phi s_t + \lambda(s_t)v_{t+1}, \]  

(14)

where \( \lambda(s_t) \) is the sensitivity function which is chosen such that the risk free rate is constant, see Campbell and Cochrane (1999) for further details.\(^{14}\) Dividend growth in the model is given by:

\[ \Delta d_{t+1} = g + w_{t+1}. \]  

(15)

We use the same calibrated monthly parameters as in Campbell and Cochrane (1999) and set the correlation between the shocks \( v_{t+1} \) and \( w_{t+1} \) equal to 0.2. We solve the model using the solution method described in Wachter (2005). Let \( D_t^{(n)} \) denote the price of a dividend at time \( t \) that is paid \( n \) periods in the future. Let \( D_{t+1} \) denote the realized dividend in period \( t + 1 \). The price of the first dividend strip is simply given by:

\[ D_t^{(1)} = E_t (M_{t+1}D_{t+1}) = D_t E_t \left( M_{t+1} \frac{D_{t+1}}{D_t} \right). \]  

(16)

The following recursion then allows us to compute the remaining dividend strips:

\[ D_t^{(n)} = E_t (M_{t+1}D_{t+1}^{n-1}) \]  

(17)

The return on the \( n^{th} \) dividend strip is given by:

\[ R_{n,t+1} = \frac{D_t^{(n-1)}}{D_t^{(n)}} \]  

(18)

We simulate from the model and compute for each dividend claim with a maturity of \( n \) months the average annualized excess return (risk premium), \( E(R_{n,t+1}) - R_f \), the annualized volatility \( \sigma(R_{n,t+1}) \), and the Sharpe ratio. The results are plotted in Figure 12 for the first 480 months (40 years). The graph shows that the term structure of expected

\(^{14}\)Wachter (2006) considers an extension to also match the term structure of interest rates.
returns and volatilities is upward sloping and the Sharpe ratio is upward sloping as well. The early dividend strips have a low average excess return equal to 1%.

### 5.2 Long-run risks

We then consider a long run risk model. We use the model and monthly calibration by Bansal and Shaliastovich (2009) which is designed to match return moments across stock, bond and foreign exchange markets. However, highly comparable results are achieved by using the model and calibration by either Bansal and Yaron (2004) or Bansal, Kiku, and Yaron (2006). The log stochastic discount factor in this model is given by:

\[
m_{t+1} = \mu_s + s_x x_t + s_g \left( \sigma^2_{gt} - \sigma^2_g \right) + s_s \left( \sigma^2_{xt} - \sigma^2_x \right) - \lambda \eta \sigma_g \eta_{t+1} - \lambda \sigma_x \eta_{t+1} - \lambda \sigma_g \sigma_w w_{g,t+1} - \lambda \sigma_x \sigma_{x,w} w_{x,t+1}.
\]

The processes for consumption growth \( \Delta c_{t+1} \), dividend growth \( \Delta d_{t+1} \) are given by:

\[
\begin{align*}
\Delta c_{t+1} &= x_t + \mu_g + \sigma_{gt} \eta_{t+1}, \\
\Delta d_{t+1} &= \mu_d + \phi_x x_t + \varphi_d \sigma_{gt} \eta_{d,t+1}
\end{align*}
\]

The three state variables in the model are \( x_t \), which is the slowly time-varying mean of consumption and dividend growth (the long run risk component), \( \sigma^2_{xt} \), which is the stochastic variance of the long-run risk component, and \( \sigma^2_{gt} \), which is the stochastic variance of the short-term risk component.

\[
\begin{align*}
x_{t+1} &= \rho x_t + \sigma_{xt} \eta_{t+1}, \\
\sigma^2_{g,t+1} &= \sigma^2_g + \nu_g \left( \sigma^2_{gt} - \sigma^2_g \right) + \sigma_{gw} w_{g,t+1}, \\
\sigma^2_{x,t+1} &= \sigma^2_x + \nu_x \left( \sigma^2_{xt} - \sigma^2_x \right) + \sigma_{xw} w_{x,t+1}.
\end{align*}
\]

We compute dividend strips in the same manner as described in the previous subsection, and we compute the average annualized excess return, volatility and Sharpe ratio. More details on how to compute the dividend strips are provided in Appendix B. The results are plotted in Figure 13. Interestingly, the results are very similar to the habit formation model. The terms structure of expected returns and volatilities is upward sloping and the Sharpe ratio is upward sloping as well.
5.3 Variable rare disasters

We then consider the variable rare disasters model by Gabaix (2009). In this case, the stochastic discount factor is given by:

\[
\frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 1 & \text{if there is no disaster at time } t+1 \\ B_{t+1}^{-\gamma} & \text{if there is a disaster at time } t+1 \end{cases}
\]  

(19)

where \(B_{t+1}\) measures the drop in consumption in case a disaster hits, and \(\delta\) is the sum of the subjective discount factor and the aggregate growth rate. The dividend process for stock \(i\) takes the form:

\[
\frac{D_{i,t+1}}{D_{it}} = e^{g_{iD}} \left(1 + \varepsilon^D_{i,t+1}\right) \times \begin{cases} 1 & \text{if there is no disaster at time } t+1 \\ F_{i,t+1} & \text{if there is a disaster at time } t+1 \end{cases}
\]  

(20)

where \(\varepsilon^D_{i,t+1} > -1\) is an independent shock with mean 0 and variance \(\sigma^2_D\), and \(F_{i,t+1} > 0\) is the recovery rate in case a disaster happens. The resilience of asset \(i\) is defined as:

\[
H_{it} = p_t E^D_{it} \left[B_{t+1}^{-\gamma}F_{i,t+1} - 1\right]
\]

where the superscript \(D\) signifies conditioning on the disaster event and \(p_t\) is the probability of a disaster. Instead of modeling each component of \(H_{it}\), Gabaix (2009) assumes that \(\hat{H}_{it} \equiv H_{it} - H_{i^*}\), follows a near-AR(1) process given by:

\[
\hat{H}_{i,t+1} = \frac{1 + H_{i^*}}{1 + \hat{H}_{it}} e^{-\phi_H} \hat{H}_{it} + \varepsilon^H_{i,t+1}
\]  

(21)

where \(\varepsilon^H_{i,t+1}\) has a conditional mean of 0 and a variance of \(\sigma^2_H\), and \(\varepsilon^H_{i,t+1}\) and \(\varepsilon^D_{i,t+1}\) are uncorrelated with the disaster event. Further details are provided in Appendix D.

In this model, the term structure of expected returns is flat. The reason is that strips of all maturities are exposed to the same risk in case of a disaster. Further, the return volatility is increasing with maturity. The reason is that longer maturity strips have a higher volatility because their duration is higher. As a result, the Sharpe ratio is downward sloping.

5.4 Lettau and Wachter (2007)

We finally consider the model by Lettau and Wachter (2007), which is designed to generate a downward sloping term structure of expected returns. In their framework, the stochastic
discount factor, which is specified exogenously, is given by:

\[ M_{t+1} = \exp(-r_f - \frac{1}{2} x_t^2 + x_t \varepsilon_{d,t+1}) \] (22)

where \( x_t \) drives the price of risk and follows an AR(1) process:

\[ x_{t+1} = (1 - \phi_x) \bar{x} + \phi_x x_t + \sigma_x \varepsilon_{t+1} \] (23)

where \( \varepsilon_{t+1} \) is a 3x1 vector of shocks and \( \sigma_x \) is 1x3 vector. Dividend growth is predictable and given by:

\[ \Delta d_{t+1} = g + z_t + \sigma_d \varepsilon_{t+1} \] (24)

where

\[ z_{t+1} = \rho_z z_t + \sigma_z \varepsilon_{t+1} \] (25)

We use their quarterly calibration and compute dividend strips using the essentially affine structure of the setup\(^\text{15}\). For more details on the calibration and the computation of dividend strips within their model, we refer to Lettau and Wachter (2007).

As before, we report for each dividend strip \( n \) the average annualized excess return (risk premium), \( E(R_{n,t+1}) - R_f \), the annualized volatility \( \sigma(R_{n,t+1}) \) and the Sharpe ratio. The results are plotted in Figure\(^\text{14}\). The term structure for the risk premium is downward sloping and the term structure of volatilities is initially upward sloping up until 8 years, and downward sloping thereafter. The Sharpe ratio is downward sloping.

6 Robustness

In this section, we perform several robustness checks of our empirical results.

6.1 Alternative selection criteria

In constructing the prices of dividend strips, we take the median across all matches of put and call contracts with the same maturity and strike price, for a given maturity and for which the prices are quoted within the same second. We select the time frame from 10am to 2pm. We now consider six alternative procedures to construct dividend prices. In all cases, we report the summary statistics of dividend returns for strategy 1, and

\(^{15}\)We apply a similar method to compute the dividend strips in the long run risk model as described in appendix B.
the CAPM alpha and beta.\footnote{The results for dividend steepener are highly comparable and are not reported for brevity.} For Alternative 1, we first minimize the time difference between contracts with the same maturity and strike price, we then select the moneyness that is closest to one for a given maturity, and if multiple matches are found, we take the median across the matches for that particular maturity. For Alternative 2, we first minimize the time difference between contracts with the same maturity and strike price, we then select the smallest bid-ask spread for a given maturity, and if multiple matches are found, we take the median across the matches for that particular maturity. In case of Alternative 3, we use the same matching procedure as in the benchmark case, but narrow the time frame to 10am to 11am, and in case of Alternative 4, we consider the time frame from 1pm to 2pm. We exclude the lunch period for the latter two alternative matching procedures, which might be a period of lower liquidity. In case of Alternative 5, we consider all matches between put and call contracts for a given maturity and strike price, but instead of minimizing over the time difference first, we take the median right away. The advantage is that we take the median across a larger set of contracts, but the time difference between the quotes might not be zero, which introduces noise. In practice, there are so many quotes that the difference time stamps of quotes is in most cases small. Finally, in case of Alternative 6, we again match all call and put contracts based on maturity and the strike price. However, instead of minimizing the time difference first, we first minimize over the bid-ask spread, and for the set of matches with the same spread for a given maturity, we take the minimum time difference. If multiple matches exist for a particular maturity, we take the median across the matches that have the smallest bid-ask spread and time difference.

The results are presented in Table 7, in which $A_i$ corresponds to Alternative $i$. Even though the numbers change slightly across different matching procedures, which is not unexpected, none of our main results is overturned for any of the cases. The dividend strategy earns high average returns, has a relatively high volatility, has a modest CAPM beta, and, as a consequence, a substantial CAPM alpha. It seems challenging to construct an argument based on microstructure issues that explains all seven empirical facts of dividend strategies, and is robust to all seven matching procedures we consider.

### 6.2 Dividend prices implied by futures contracts

As an alternative robustness check, we consider a different market to synthetically construct dividend prices. Instead of relying on options markets, we use data on index futures. As discussed above, index futures do not have as long maturities as index
options, but we have access to maturities up to one year. Figure 11 displays the dividend prices for a 6-month and 1-year contract implied by either futures data or options data. To make both series stationary, we scale the price series by the level of the S&P500 index. The relative price series clearly have the same level and are highly correlated; the full-sample correlation equals 94% for the 6-month contract and 91% for the 1-year contract. As such, explanations of our findings must also be able to explain the same phenomenon in futures markets. Explanations for all facts solely based on market microstructure are therefore, in our view, less convincing as index futures markets are among the most liquid asset markets available.

6.3 Sensitivity to interest rates

To explore the sensitivity to the LIBOR rates that we use, we perform the following sensitivity analysis. We recompute the dividend prices for both strategies, changing the interest rate by $\delta$, where we let $\delta$ vary between -50 and +50 basis points. This leads to the following dividend prices:

$$P_{t,T} = p_{t,T} - c_{t,T} + S_t - X e^{-(r_{t,T}+\delta)(T-t)},$$

(26)

In case of the dividend steepener, we compute:

$$P_{t,T_1,T_2} = P_{t,T_2} - P_{t,T_1} = p_{t,T_2} - p_{t,T_1} - c_{t,T_2} + c_{t,T_1} - X \left( e^{-(r_{t,T_2}+\delta)(T_2-t)} - e^{-(r_{t,T_1}+\delta)(T_1-t)} \right).$$

(27)

For each value of $\delta$, we recompute the dividend returns and compute the time-series average of $R_{1,t+1}$ and $R_{2,t+1}$. This allows us to assess what increase in interest rates we would need to drive the average return on both strategies to zero. The results are summarized in Table 8. We find that the average return for strategy 1 ($E[R_{1,t+1}]$) is zero when $\delta = +50bp$. Regardless of whether one considers this number to be small or large, more importantly this interest rate increase does not drive the average return on the dividend steepener to zero. As $\delta$ enters twice in Equation (27) the two terms offset each other. As such, within the range of values that we consider, there is no value for $\delta$ for which the average return on the dividend steepener can be driven to zero. Further, the baseline case ($\delta = 0bp$) appears to provide a lower bound on the expected return on the dividend steepener.
7 Further applications

In this section, we illustrate two other applications that can be explored using the dividend strips we compute in this paper.

7.1 Risk-neutral growth rates

European data. Japan/US

7.2 Stochastic discount factor decompositions

Building on Bansal and Lehman (1997), Hansen, Heaton, and Li (2008) and Hansen and Scheinkman (2009) show how to decompose the pricing kernel into a permanent and temporary component. These decompositions are useful for various reasons. Alvarez and Jermann (2005) for instance show that the ratio of the variance of the permanent component to the overall variance is equal to one minus the ratio of the long-term bond risk premium to the maximum risk premium across all securities. This insight can be used to identify pricing factors and to generate additional restrictions for general equilibrium asset pricing models. In addition, these decompositions are useful to understand how future dividend prices respond to a shock to a macro-economic state variable today, see Borovicka, Hansen, Hendricks, and Scheinkman (2009). Borovicka, Hansen, Hendricks, and Scheinkman (2009) largely use these results to point out differences across asset pricing models, but there is no empirical counterpart yet to which this models can be compared. The methods we develop in this paper might be useful to advance our understanding of the decomposition of the stochastic discount factor.

7.3 Market-implied expected returns and expected growth rates

Binsbergen and Kojien (2010) show how to use filtering methods to estimate expected returns and expected growth rates. Filtering methods are required as the price-dividend ratio is an affine function of expected returns and expected growth rates (see also Section 5), which are both latent. However, if we use exactly the same model to price dividend strips, it follows immediately that all dividend strips are affine in the same two state variables, but with different loadings. Assuming that the model is correctly specified, this implies, reminiscent to the term structure literature, that we can invert any

\[^{17}\text{See for instance Kojien, Lustig, and Van Nieuwerburgh (2009) and Kojien, Lustig, Van Nieuwerburgh, and Verdelhan (2010).}\]
two dividend strips to recover market-implied expected returns and growth rates.

8 Conclusion

We study the pricing of short-term assets whose payoff equals the dividends of the aggregate stock market during a period of up to three years. To compute these prices, we apply the put-call parity to a new data set of liquid, exchange-traded S&P500 options. We compare the asset pricing properties of the claim to short-term dividends to the pricing of the aggregate stock market, which is the claim to all future dividends. Using this approach, we find that the short-term asset has a high expected returns, a beta to the market of 0.5, is excessively volatile, and has returns that are highly predictable. The returns on short-term dividend claims cannot be explained by standard asset pricing models, which makes such claims important candidate test assets. We compare our empirical results to their theoretical equivalents in leading asset pricing models and find that none of them predict the empirical findings we document.

Additional notes are available upon request.
References


A Details dividend returns

The two trading strategies described in Section 3.3 can be implemented for different maturities $T$. The specific maturities we follow for trading strategy 1 vary between 1.9 years and 1.3 years. To be precise, for trading strategy 1, we go long in the 1.874 year dividend claim on January 31st 1996, collect the dividend during February and sell the claim on February 29th 1996 to compute the return. The claim then has a remaining maturity of 1.797 years. We buy back the claim (or alternatively, we never sold it), go long in the 1.797 year claim, collect the dividend, and sell it on March 29th 1996. We follow this strategy until July 31st 1996 at which time the remaining maturity is 1.381 years. On this date a new 1.881 year contract is available so we restart the investment cycle at this time, and continue until May of 2009, which is the end of our sample.

For trading strategy 2, we follow the same maturities, apart from the fact that we go long in the 1.874 year dividend claim and short in the 0.874 dividend claim on January 31st 1996. On July 31st 1996 the remaining maturities are 1.381 years and 0.381 years at which point we restart the investment cycle in the 1.881 year contract and the 0.881 year contract available at that time.

B Forecasting returns and dividend growth rates

We follow Binsbergen and Koijen (2010) and use filtering techniques to predict future dividend growth rates and returns. Let $r_{t+1}$ denote the total log return on the index:

$$r_{t+1} \equiv \log \left( \frac{S_{t+1} + D_{t+1}}{S_t} \right),$$

where let $PD_t$ denote the price-dividend ratio:

$$PD_t \equiv \frac{S_t}{D_t};$$

and let $\Delta d_{t+1}$ denote the aggregate log dividend growth rate:

$$\Delta d_{t+1} \equiv \log \left( \frac{D_{t+1}}{D_t} \right).$$
We model both expected returns ($\mu_t$) and expected dividend growth rates ($g_t$) as an AR(1)-process:

\begin{align}
\mu_{t+1} &= \delta_0 + \delta_1 (\mu_t - \delta_0) + \epsilon^\mu_{t+1}, \\
g_{t+1} &= \gamma_0 + \gamma_1 (g_t - \gamma_0) + \epsilon^g_{t+1},
\end{align}

(29)  

(30)

where $\mu_t \equiv E_t [r_{t+1}]$ and $g_t \equiv E_t [\Delta d_{t+1}]$. The distribution of the shocks $\epsilon^\mu_{t+1}$ and $\epsilon^g_{t+1}$ is specified below. The realized dividend growth rate is equal to the expected dividend growth rate plus an orthogonal shock:

$$
\Delta d_{t+1} = g_t + \epsilon^D_{t+1}.
$$

Defining $pd_t \equiv \log (PD_t)$, we can write the log-linearized return as:

$$
r_{t+1} \simeq \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t,
$$

with $pd = E[pd_t]$, $\kappa = \log \left(1 + \exp \left(\bar{pd}\right)\right) - \rho \bar{pd}$ and $\rho = \frac{\exp(\bar{pd})}{1+\exp(\bar{pd})}$, as in Campbell and Shiller (1988). If we iterate this equation, and using the AR(1) assumptions (29)-(30), it follows that:

$$
\begin{align}
\bar{pd}_t &= A - B_1 (\mu_t - \delta_0) + B_2 (g_t - \gamma_0),
\end{align}
$$

with $A = \frac{\kappa}{1-\rho} + \frac{\gamma_0 - \delta_0}{1-\rho\gamma_1}$, $B_1 = \frac{1}{1-\rho_1}$, and $B_2 = \frac{1}{1-\rho_1}$. The log price-dividend ratio is linear in the expected return $\mu_t$ and the expected dividend growth rate $g_t$. The loading of the price-dividend ratio on expected returns and expected dividend growth rates depends on the relative persistence of these variables ($\delta_1$ versus $\gamma_1$). The three shocks in the model, which are shocks to expected dividend growth rates ($\epsilon^g_{t+1}$), shocks to expected returns ($\epsilon^\mu_{t+1}$), and realized dividend growth shocks ($\epsilon^d_{t+1}$), have mean zero, covariance matrix

$$
\Sigma \equiv \text{var} \begin{pmatrix}
\epsilon^g_{t+1} \\
\epsilon^\mu_{t+1} \\
\epsilon^d_{t+1}
\end{pmatrix} = \begin{bmatrix}
\sigma^2_g & \sigma_{g\mu} & \sigma_{gD} \\
\sigma_{g\mu} & \sigma^2_\mu & \sigma_{\muD} \\
\sigma_{gD} & \sigma_{\muD} & \sigma^2_D
\end{bmatrix},
$$

and are independent and identically distributed over time. Further, in the maximum likelihood estimation procedure, we assume that the shocks are jointly normally distributed.

We subsequently perform unconditional maximum likelihood estimation to obtain
estimates for all the parameters and obtain filtered series for \( \mu_t \) and \( g_t \). The \( R^2 \) values are computed as:

\[
R_{\text{ret}}^2 = 1 - \frac{\text{var} \left( r_{t+1} - \mu_t^F \right)}{\text{var} \left( r_t \right)},
\]

\[
R_{\text{Div}}^2 = 1 - \frac{\text{var} \left( \Delta d_{t+1} - g_t^F \right)}{\text{var} \left( \Delta d_{t+1} \right)},
\]

where \( \text{var} \) is the sample variance, \( \mu_t^F \) is the filtered series for expected returns (\( \mu_t \)) and \( g_t^F \) is the filtered series for expected dividend growth rates (\( g_t \)).

C Dividend strips in the long-run risks model

In this appendix we derive dividend strips in a long run risk model as calibrated by Bansal and Shaliastovich (2009). In their model, the log stochastic discount factor is given by:

\[
m_{t+1} = \mu_s + s_x x_t + s_{gs} (\sigma_{gt}^2 - \sigma_g^2) + s_{xs} (\sigma_{xt}^2 - \sigma_x^2) - \lambda_\eta \sigma_{gt} \eta_{t+1} - \lambda_e \sigma_{xt} e_{t+1} - \lambda_{gw} \sigma_{gw} w_{g,t+1} - \lambda_{xw} \sigma_{xw} w_{x,t+1}.
\]

The short rate follows from:

\[
E_t (\exp (m_{t+1})) = \exp \left( E_t (m_{t+1}) + \frac{1}{2} \text{var}_t (m_{t+1}) \right)
\]

\[
= \exp \left( \mu_s + s_x x_t + s_{gs} (\sigma_{gt}^2 - \sigma_g^2) + s_{xs} (\sigma_{xt}^2 - \sigma_x^2) + \frac{1}{2} \left[ \lambda_\eta^2 \sigma_{gt}^2 + \lambda_e^2 \sigma_{xt}^2 + \lambda_{gw}^2 \sigma_{gw}^2 + \lambda_{xw}^2 \sigma_{xw}^2 \right] \right)
\]

\[
= \exp \left( w_0 + w_1 x_t + w_2 \sigma_{gt}^2 + w_3 \sigma_{xt}^2 \right)
\]

with:

\[
w_0 = \mu_s - s_{gs} \sigma_g^2 - s_{xs} \sigma_x^2 + \frac{1}{2} \left[ \lambda_{gw}^2 \sigma_{gw}^2 + \lambda_{xw}^2 \sigma_{xw}^2 \right],
\]

\[
w_1 = s_x,
\]

\[
w_2 = s_{gs} + \frac{1}{2} \lambda_\eta^2,
\]

\[
w_3 = s_{xs} + \frac{1}{2} \lambda_e^2.
\]
The state variables satisfy the following dynamics:

\[ \Delta c_{t+1} = x_t + \mu_g + \sigma_g \eta_{t+1}, \]
\[ x_{t+1} = \rho x_t + \sigma_x \epsilon_{t+1}, \]
\[ \sigma_{g,t+1}^2 = \sigma_g^2 + \nu_g \left( \sigma_{g,t}^2 - \sigma_g^2 \right) + \sigma_{gw} w_{g,t+1}, \]
\[ \sigma_{x,t+1}^2 = \sigma_x^2 + \nu_x \left( \sigma_{x,t}^2 - \sigma_x^2 \right) + \sigma_{xw} w_{x,t+1}, \]
\[ \Delta d_{t+1} = \mu_d + \phi_x x_t + \phi_d \sigma_g \eta_{d,t+1}. \]

All shocks are independent, apart from: \( \tau_{g,t} = corr(\eta_{t+1}, \eta_{d,t+1}) = cov(\eta_{t+1}, \eta_{d,t+1}) \). All unknown coefficients are defined in Koijen, Lustig, VanNieuwerburgh, and Verdelhan (2009). The 1-period dividend strip follows from:

\[ D_{t}^{(1)} = D_t E_t \left( M_{t+1} \frac{D_{t+1}}{D_t} \right) \]
\[ = D_t E_t \left( \exp \left( \mu_s + s_x x_t + s_g \left( \sigma_{g,t}^2 - \sigma_g^2 \right) + s_{gs} \sigma_{g,t}^2 + s_{gx} \sigma_x^2 - \sigma_x^2 \right) \right) \]
\[ + \mu_d + \phi_x x_t + \phi_d \sigma_g \eta_{d,t+1} \]
\[ = \exp \left( \mu_s + s_x x_t + s_g \left( \sigma_{g,t}^2 - \sigma_g^2 \right) + s_{gs} \sigma_{g,t}^2 + s_{gx} \sigma_x^2 - \sigma_x^2 + \mu_d + \phi_x x_t \right) \times \]
\[ \exp \left( \frac{1}{2} \left[ \lambda^2 _g \sigma_{g,t}^2 + \lambda^2 _x \sigma_x^2 + \lambda^2 _g \sigma_{gw}^2 + \lambda^2 _x \sigma_{xw}^2 \right] - \lambda \eta \phi_{gd} \right) \]
\[ = D_t \exp \left( H_0^{(1)} + H_1^{(1)} x_t + H_2^{(1)} \sigma_{g,t}^2 + H_3^{(1)} \sigma_x^2 \right). \]

This leads to:

\[ H_0^{(1)} = \mu_s - s_g \sigma_g^2 - s_{gs} \sigma_{g,t}^2 + \mu_d + \frac{1}{2} \left[ \lambda^2 _g \sigma_{gw}^2 + \lambda^2 _x \sigma_{xw}^2 \right], \]
\[ H_1^{(1)} = s_x + \phi_x, \]
\[ H_2^{(1)} = s_{gs} + \frac{1}{2} \lambda^2 _g - \lambda \eta \phi_{gd}, \]
\[ H_3^{(1)} = s_{gs} + \frac{1}{2} \lambda^2 _x. \]
As before, let $D_t^{(n)}$ denote the price of a dividend at time $t$ that is paid out in $n$ period. The following relationship then holds:

$$D_t^{(n)} = E_t \left( D_{t+1}^{(n-1)} M_{t+1} \right)$$

$$= D_t E_t \left( \exp \left( H_0^{(n-1)} + H_1^{(n-1)} x_{t+1} + H_2^{(n-1)} \sigma_{gt, t+1}^2 H_3^{(n)} \sigma_{zt, t+1}^2 \right) M_{t+1} \frac{D_{t+1}}{D_t} \right)$$

$$= D_t E_t \left( \exp \left( H_0^{(n-1)} + H_1^{(n-1)} x_{t+1} + H_2^{(n-1)} \sigma_{gt, t+1}^2 H_3^{(n)} \sigma_{zt, t+1}^2 \right) + \mu_s + s_x x_t + s_{gs} \left( \sigma_{gt}^2 - \sigma_g^2 \right) + s_{xs} \left( \sigma_{zt}^2 - \sigma_x^2 \right) - \lambda_x \sigma_{zt} \eta_{t+1} + \lambda_g \sigma_{gw} \sigma_{gw, t+1} - \lambda_{xw} \sigma_{xw} \sigma_{xw, t+1} + \mu_d + \phi_x x_t + \varphi_d \sigma_{gt} \eta_{d, t+1} \right)$$

which can be rewritten as:

$$D_t^{(n)} = D_t E_t \left( \exp \left( H_0^{(n-1)} + H_1^{(n-1)} \rho x_t + H_2^{(n-1)} \left( \sigma_g^2 + \nu_g \left( \sigma_{gt}^2 - \sigma_g^2 \right) \right) \right) + H_3^{(n-1)} \left( \sigma_x^2 + \nu_x \left( \sigma_{zt}^2 - \sigma_x^2 \right) \right) + \mu_s + s_x x_t + s_{gs} \left( \sigma_{gt}^2 - \sigma_g^2 \right) + s_{xs} \left( \sigma_{zt}^2 - \sigma_x^2 \right) \right)$$

$$- \lambda_x \sigma_{zt} \eta_{t+1} + \left( H_1^{(n-1)} - \lambda_g \sigma_{gw} \sigma_{gw, t+1} - \lambda_{xw} \sigma_{xw} \sigma_{xw, t+1} + \mu_d + \phi_x x_t + \varphi_d \sigma_{gt} \eta_{d, t+1} \right)$$

$$= D_t \exp \left( H_0^{(n)} + H_1^{(n)} x_t + H_2^{(n)} \sigma_{gt}^2 + H_3^{(n)} \sigma_{zt}^2 \right) .$$

This implies:

$$H_0^{(n)} = H_0^{(n-1)} + H_2^{(n-1)} \left( 1 - \nu_g \right) \sigma_g^2 + H_3^{(n-1)} \left( 1 - \nu_x \right) \sigma_x^2$$

$$+ \mu_s - s_{gs} \sigma_g^2 - s_{xs} \sigma_x^2 + \mu_d$$

$$+ \frac{1}{2} \left[ \left( H_2^{(n-1)} - \lambda_g \sigma_{gw} \right)^2 \sigma_{gw}^2 + \left( H_3^{(n-1)} - \lambda_{xw} \right)^2 \sigma_{xw}^2 \right] ,$$

$$H_1^{(n)} = H_1^{(n-1)} \rho + s_x + \phi_x ,$$

$$H_2^{(n)} = H_2^{(n-1)} \nu_g + s_{gs} \sigma_g + \lambda_g \sigma_{gt} \sigma_{gt} + \frac{1}{2} \lambda_x \sigma_{gt}^2 \tau_g \sigma_{gt} ,$$

$$H_3^{(n)} = H_3^{(n-1)} \nu_x + s_{xs} \sigma_x + \frac{1}{2} \left( H_1^{(n-1)} - \lambda_e \right) .$$
D Dividend strips in the rare disasters model

The setup of the Barro-Rietz rare disasters model as presented by Gabaix (2009) is as follows. Let there be a representative agent with utility given by:

\[
E_0 \left[ \sum_{t=0}^{\infty} e^{-\rho t} \frac{C_{i,t}^{1-\gamma}}{1-\gamma} \right]
\]

(33)

At each period consumption growth is given by:

\[
\frac{C_{t+1}}{C_t} = e^g \times \begin{cases} 
1 & \text{if there is no disaster at time } t+1 \\
B_{t+1} & \text{if there is a disaster at time } t+1 
\end{cases}
\]

(34)

The pricing kernel is then given by:

\[
\frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 
1 & \text{if there is no disaster at time } t+1 \\
B_{t+1}^{-\gamma} & \text{if there is a disaster at time } t+1 
\end{cases}
\]

(35)

where \(\delta = \rho + g\). The dividend process for stock \(i\) takes the form:

\[
\frac{D_{i,t+1}}{D_{it}} = e^{g_{i,D}} (1 + \varepsilon_{i,t+1}^D) \times \begin{cases} 
1 & \text{if there is no disaster at time } t+1 \\
F_{i,t+1} & \text{if there is a disaster at time } t+1 
\end{cases}
\]

(36)

where \(\varepsilon_{i,t+1}^D > -1\) is an independent shock with mean 0 and variance \(\sigma_D^2\), and \(F_{i,t+1} > 0\) is the recovery rate in case a disaster happens. The resilience of asset \(i\) is defined as:

\[
H_{it} = p_i E_i^D [B_{t+1}^{-\gamma} F_{i,t+1} - 1]
\]

where the superscript \(D\) signifies conditioning on the disaster event. Define \(\hat{H}_{it} = H_{it} - H_{i^*}\), which follows a near-AR(1) process given by:

\[
\hat{H}_{i,t+1} = \frac{1 + H_{i^*}}{1 + H_{it}} e^{-\delta_H} \hat{H}_{it} + \varepsilon_{i,t+1}^H
\]

where \(\varepsilon_{i,t+1}^H\) has a conditional mean of 0 and a variance of \(\sigma_H^2\), and \(\varepsilon_{i,t+1}^H\) and \(\varepsilon_{i,t+1}^D\) are uncorrelated with the disaster event. Under the assumptions above, the stock price is given by:

\[
P_{it} = \frac{D_{it}}{1 - e^{-\delta_i}} \left( 1 + \frac{e^{-\delta_i - h_{i^*}} \hat{H}_{it}}{1 - e^{\delta_i - \delta_H}} \right)
\]
where

\[ \delta_i = \delta - g_{iD} - h_{is} \]

\[ h_{is} = \ln H_{is} \]

Gabaix (2009) shows that the price at time \( t \) of a dividend paid in \( n \) periods is given by:

\[ D_{it}^{(n)} = D_{it} e^{-\delta_i T} \left( 1 + \frac{1 - e^{\phi H}}{\phi H} H_{it} \right) \]

and that the expected return on the strip, conditioning on no disaster is given by:

\[ E_t \left[ \ln R_{n,t+1} \right] = E_t \left[ \ln \frac{D_{t+1}^{(n-1)}}{D_{t}^{(n)}} \right] \approx \delta - H_{it} \]

The expected return is the same across maturities, because strips of all maturities are exposed to the same risk in a disaster.\(^{19}\)

The volatility of the linearized return is given by:

\[ \sigma_{n,t} = \sqrt{\sigma_D^2 + \left( \frac{1 - e^{\phi H}}{\phi H} \right)^2 \sigma_H^2} \]

which is increasing with maturity, due to the fact that higher duration cash flows are more exposed to discount rate shocks than short duration cash flows. Given that the expected return is constant across maturities and the volatility is increasing with maturity, the Sharpe ratio is decreasing with maturity.

\(^{19}\)We thank Xavier Gabaix for providing us with this derivation.
## Table 1: Descriptive Statistics

The table presents descriptive statistics of the monthly returns on the two trading strategies described in the main text. As the volatility in the second half of the sample is lower than in the first half of the sample, we also present descriptive statistics for two subsamples: 1996:2-2002:12 and 2003:1-2009:5.

<table>
<thead>
<tr>
<th></th>
<th>$R_{1,t}$</th>
<th>$R_{2,t}$</th>
<th>Market</th>
<th>S&amp;P500</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Full sample 1996:2 - 2009:5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0120</td>
<td>0.0115</td>
<td>0.0054</td>
<td>0.0049</td>
</tr>
<tr>
<td>Median</td>
<td>0.0097</td>
<td>0.0148</td>
<td>0.0130</td>
<td>0.0104</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0791</td>
<td>0.0979</td>
<td>0.0492</td>
<td>0.0472</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.1155</td>
<td>0.0876</td>
<td>0.0514</td>
<td>0.0426</td>
</tr>
<tr>
<td>Observations</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>160</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: First half sample 1996:2 - 2002:12</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0159</td>
<td>0.0139</td>
<td>0.0060</td>
<td>0.0065</td>
</tr>
<tr>
<td>Median</td>
<td>0.0171</td>
<td>0.0231</td>
<td>0.0136</td>
<td>0.0093</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0986</td>
<td>0.1212</td>
<td>0.0528</td>
<td>0.0514</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.1242</td>
<td>0.0843</td>
<td>0.0564</td>
<td>0.0456</td>
</tr>
<tr>
<td>Observations</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: Second half sample 2003:1 - 2009:5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0079</td>
<td>0.0089</td>
<td>0.0047</td>
<td>0.0031</td>
</tr>
<tr>
<td>Median</td>
<td>0.0077</td>
<td>0.0067</td>
<td>0.0129</td>
<td>0.0112</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0508</td>
<td>0.0646</td>
<td>0.0453</td>
<td>0.0424</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.1139</td>
<td>0.1050</td>
<td>0.0587</td>
<td>0.0248</td>
</tr>
<tr>
<td>Observations</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
</tr>
</tbody>
</table>
### Table 2: Monthly returns on the two trading strategies and the market portfolio.
The table presents OLS regressions of the returns on trading strategies 1 and 2 (dependent variables) on the market portfolio. Newey-West standard errors in parentheses. When an AR(1) term is included, the intercept is adjusted with the AR(1) coefficient, such that the intercept is comparable to the regressions without AR(1) term.

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>$R_{1,t+1} - R_{f,t}$</th>
<th>$R_{2,t+1} - R_{f,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>0.0079 0.0077</td>
<td>0.0073 0.0069</td>
</tr>
<tr>
<td></td>
<td>(0.0052) (0.0048)</td>
<td>(0.0054) (0.0049)</td>
</tr>
<tr>
<td>mktrf</td>
<td>0.4879 0.5199</td>
<td>0.4912 0.5374</td>
</tr>
<tr>
<td></td>
<td>(0.1642) (0.1543)</td>
<td>(0.1895) (0.1680)</td>
</tr>
<tr>
<td>AR(1)</td>
<td>- -0.2935</td>
<td>- -0.3396</td>
</tr>
<tr>
<td></td>
<td>- (0.1088)</td>
<td>- (0.0827)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0914 0.1764</td>
<td>0.0606 0.1727</td>
</tr>
</tbody>
</table>

### Table 3: Monthly Returns on the Two Trading Strategies and the Three Factor Model.
The table presents OLS regressions of the returns on trading strategies 1 and 2 (dependent variables) on the Fama French three factor model. Newey-West standard errors in parentheses. When an AR(1) term is included, the intercept is adjusted with the AR(1) coefficient, so that the intercept is comparable to the regressions without AR(1) term.

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>$R_{1,t+1} - R_{f,t}$</th>
<th>$R_{2,t+1} - R_{f,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>c</td>
<td>0.0069 0.0070</td>
<td>0.0057 0.0056</td>
</tr>
<tr>
<td></td>
<td>(0.0045) (0.0043)</td>
<td>(0.0049) (0.0045)</td>
</tr>
<tr>
<td>mktrf</td>
<td>0.5176 0.5426</td>
<td>0.5987 0.6248</td>
</tr>
<tr>
<td></td>
<td>(0.1461) (0.1368)</td>
<td>(0.1587) (0.1376)</td>
</tr>
<tr>
<td>hml</td>
<td>0.1954 0.1765</td>
<td>0.4305 0.3988</td>
</tr>
<tr>
<td></td>
<td>(0.1965) (0.1862)</td>
<td>(0.2283) (0.2488)</td>
</tr>
<tr>
<td>smb</td>
<td>0.0959 0.1074</td>
<td>-0.0158 0.043</td>
</tr>
<tr>
<td></td>
<td>(0.1541) (0.1519)</td>
<td>(0.1749) (0.1623)</td>
</tr>
<tr>
<td>AR(1)</td>
<td>- -0.2917</td>
<td>- -0.3306</td>
</tr>
<tr>
<td></td>
<td>- (0.1094)</td>
<td>- (0.0822)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0982 0.1823</td>
<td>0.0854 0.1919</td>
</tr>
</tbody>
</table>
Table 4: Estimates of the GARCH(1,1) model
The top panel provides the estimates of the mean equation; the bottom panel displays the estimates of the variance model. The first two columns report the results for the dividend return strategies, and the third column provides the results for the S&P500.

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>$R_{1,t+1}$</th>
<th>$R_{2,t+1}$</th>
<th>$R_{SP500,t+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean equation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c$</td>
<td>0.0073</td>
<td>0.0083</td>
<td>0.0069</td>
</tr>
<tr>
<td>(0.0046)</td>
<td>(0.0041)</td>
<td>(0.0038)</td>
<td></td>
</tr>
<tr>
<td>AR(1)</td>
<td>-0.2905</td>
<td>-0.3954</td>
<td>0.0790</td>
</tr>
<tr>
<td>(0.1059)</td>
<td>(0.1094)</td>
<td>(0.0996)</td>
<td></td>
</tr>
</tbody>
</table>

| Variance equation | | | |
| $c$ | $1.3 \times 10^{-4}$ | $2.4 \times 10^{-4}$ | $6.8 \times 10^{-5}$ |
| (7.2 $\times 10^{-5}$) | (1.5 $\times 10^{-4}$) | (5.2 $\times 10^{-5}$) |
| squared residual | 0.1324 | 0.1985 | 0.1937 |
| (0.0421) | (0.0621) | (0.0761) |
| GARCH(1) | 0.8728 | 0.8174 | 0.8071 |
| (0.0275) | (0.0518) | (0.0610) |

Table 5: Maximum-likelihood estimates
We present the estimation results of the present-value model. The model is estimated by unconditional maximum likelihood using data over three different sample periods, 1946-2007, 1970-2007, 1989-2007 on cash-invested dividend growth rates and the corresponding price-dividend ratio.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_0$</td>
<td>0.0899</td>
<td>0.0782</td>
<td>0.0885</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.9267</td>
<td>0.9349</td>
<td>0.8734</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>0.0616</td>
<td>0.0564</td>
<td>0.0716</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.4856</td>
<td>0.6966</td>
<td>0.7901</td>
</tr>
<tr>
<td>$\sigma_\mu$</td>
<td>0.0179</td>
<td>0.0166</td>
<td>0.0318</td>
</tr>
<tr>
<td>$\sigma_g$</td>
<td>0.0465</td>
<td>0.0328</td>
<td>0.0324</td>
</tr>
<tr>
<td>$\sigma_D$</td>
<td>0.0035</td>
<td>0.0038</td>
<td>0.0025</td>
</tr>
<tr>
<td>$\sigma_{g\mu}$</td>
<td>0.4939</td>
<td>0.4981</td>
<td>0.8153</td>
</tr>
<tr>
<td>$\sigma_{\mu D}$</td>
<td>0.8581</td>
<td>0.8631</td>
<td>-0.5745</td>
</tr>
<tr>
<td>$R^2_{Ret}$</td>
<td>0.0977</td>
<td>0.0904</td>
<td>0.2451</td>
</tr>
<tr>
<td>$R^2_{Div}$</td>
<td>0.2423</td>
<td>0.4086</td>
<td>0.5670</td>
</tr>
</tbody>
</table>
Table 6: Return predictability
The table presents regressions of the monthly return series on trading strategy 1, $R_{t+1}$, on the ratio of the one-year dividend strip at time $t$, denoted by $P_{t,T}$, and the aggregated dividend paid out over the previous twelve months, where dividends are reinvested in the risk free rate. We also regress returns on a smoothed version of $P_{t,T}/D_t$, where the smoothed ratio is computed by taking a rolling average over the past three values of $P_{t,T}/D_t$.

<table>
<thead>
<tr>
<th>Dep. Var.</th>
<th>$R_{t+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C$</td>
<td>0.1904</td>
</tr>
<tr>
<td></td>
<td>(0.0452)</td>
</tr>
<tr>
<td>$P_{t,T}/D_t$</td>
<td>-0.1820</td>
</tr>
<tr>
<td></td>
<td>(0.0456)</td>
</tr>
<tr>
<td>$P_{t,T}/D_t$ smoothed</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$R^2$ 0.0983 0.0547

Table 7: Alternative selection criteria
The table presents the summary statistics of dividend strategy 1 for six alternative selection criteria (A1 to A6), which are described in the main text. The table also reports the CAPM alpha and beta.

<table>
<thead>
<tr>
<th></th>
<th>Benchmark</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0120</td>
<td>0.0124</td>
<td>0.0149</td>
<td>0.0122</td>
<td>0.0120</td>
<td>0.0121</td>
<td>0.0141</td>
</tr>
<tr>
<td>Median</td>
<td>0.0097</td>
<td>0.0139</td>
<td>0.0056</td>
<td>0.0102</td>
<td>0.0091</td>
<td>0.0091</td>
<td>0.0045</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>0.0791</td>
<td>0.1041</td>
<td>0.1332</td>
<td>0.0799</td>
<td>0.0792</td>
<td>0.0790</td>
<td>0.1291</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.1155</td>
<td>0.0913</td>
<td>0.0901</td>
<td>0.1164</td>
<td>0.1149</td>
<td>0.1165</td>
<td>0.0868</td>
</tr>
<tr>
<td>CAPM Alpha</td>
<td>0.0079</td>
<td>0.0083</td>
<td>0.0100</td>
<td>0.0081</td>
<td>0.008</td>
<td>0.0079</td>
<td>0.0094</td>
</tr>
<tr>
<td>CAPM Beta</td>
<td>0.4879</td>
<td>0.4528</td>
<td>0.8020</td>
<td>0.4844</td>
<td>0.4236</td>
<td>0.4907</td>
<td>0.7228</td>
</tr>
</tbody>
</table>
Table 8: Sensitivity to interest rates

The table presents the sensitivity analysis to interest rates. We add a constant $\delta$ to our LIBOR interest rates and recompute the sample averages of the returns for values of $\delta$ varying between -50bp and +50bp. For $\delta = 0$, our baseline results are obtained (in bold).

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>-50bp</th>
<th>-40bp</th>
<th>-20bp</th>
<th>-10bp</th>
<th>0bp</th>
<th>10bp</th>
<th>20bp</th>
<th>30bp</th>
<th>40bp</th>
<th>50bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[R_{1,t+1}]$</td>
<td>0.054</td>
<td>0.038</td>
<td>0.021</td>
<td>0.016</td>
<td><strong>0.012</strong></td>
<td>0.009</td>
<td>0.006</td>
<td>0.004</td>
<td>0.002</td>
<td>0</td>
</tr>
<tr>
<td>$E[R_{2,t+1}]$</td>
<td>0.033</td>
<td>0.021</td>
<td>0.013</td>
<td>0.012</td>
<td><strong>0.012</strong></td>
<td>0.012</td>
<td>0.012</td>
<td>0.013</td>
<td>0.014</td>
<td>0.015</td>
</tr>
</tbody>
</table>
Figure 1: Maximum maturity of LEAPS
The graph displays the maximum maturity of LEAPS contracts that is available at each point of the sample. The sample period is January 1996 up to May 2009.
Figure 2: Average number of matches
The graph shows the average number of matches of put and call contracts with strike prices and maturities that coincide, and for which the quotes are provided in the same second during the last trading day of the month. We focus on contracts with a maturity between 1 and 2 years, and average the number of matches within a year. We report the natural logarithm of the number of matches. The sample period is January 1996 up to May 2009.
Figure 3: Price dynamics of the short-term assets (Cumulative)
The graph shows the prices of the first 0.5, 1, 1.5 and 2 years of dividends. The sample period is January 1996 up to May 2009.
Figure 4: Present value of dividends as a fraction of the index value (Cumulative)
The graph shows the net present value of the first 0.5, 1, 1.5 and 2 years of dividends as a fraction of the index value as computed. The sample period is January 1996 up to May 2009.
Figure 5: Monthly returns on trading strategy 1: 1996:2-2009:5: line graph.

Figure 6: Monthly returns on trading strategy 2: 1996:2-2009:5: line graph.
Figure 7: Monthly returns on trading strategy 1: 1996:2-2009:5: histogram.

Figure 8: Monthly returns on trading strategy 2: 1996:2-2009:5: histogram.
Figure 9: Prices and realizations of dividend claims: 1996:2-2009:5.

Figure 10: Volatility of dividend returns and returns on the S&P500 based on a GARCH(1,1) model.
Figure 11: Short-term asset prices implied by futures and options
The graph shows the price of the short-term assets implied by futures and option markets. The maturity of the short-term asset equals either 0.5 year or 1 year.
Figure 12: Term Structure of the Risk Premium, Volatility and Sharpe Ratio for External Habits
The graph shows the term structures of the risk premium, the volatility and the Sharpe ratio for the Campbell Cochrane (1999) habit formation model. The graph plots the first 480 months of dividend strips, which corresponds to 40 years.
Figure 13: Term Structure of the Risk Premium, Volatility and Sharpe Ratio for the Long Run Risk Model
The graph shows the term structures of the risk premium, the volatility and the Sharpe ratio for the long run risk model as calibrated by Bansal and Shaliastovich (2009). The graph plots the first 480 months of dividend strips, which corresponds to 40 years.
Figure 14: Term Structure of the Risk Premium, Volatility and Sharpe Ratio for the Lettau Wachter (2007) Model
The graph shows the term structures of the risk premium, the volatility and the Sharpe ratio for the Lettau Wachter (2007) model. The graph plots the first 120 quarters of dividend strips, which corresponds to 40 years.